



Caractérisation des coefficients de Strickler d'un fleuve par inversion probabiliste

Mathieu Couplet, Laurent Le Brusquet, Alberto Pasanisi

► To cite this version:

Mathieu Couplet, Laurent Le Brusquet, Alberto Pasanisi. Caractérisation des coefficients de Strickler d'un fleuve par inversion probabiliste. 42èmes Journées de Statistique, 2010, Marseille, France, France. inria-00494773

HAL Id: inria-00494773

<https://inria.hal.science/inria-00494773>

Submitted on 24 Jun 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

CARACTÉRISATION DES COEFFICIENTS DE STRICKLER D'UN FLEUVE PAR INVERSION PROBABILISTE

Mathieu Couplet¹, Laurent Le Brusquet² et Alberto Pasanisi¹

1. EDF R&D, Dépt. Management des Risques Industriels, 78401 Chatou, France

2. SUPELEC, Dépt. Signaux & Systèmes Électroniques, 91192 Gif-sur-Yvette, France

Résumé : La caractérisation statistique des coefficients de rugosité du lit et des berges d'un fleuve (coefficients de Strickler) permet à la fois d'évaluer les risques d'inondation et de prévoir le renforcement des ouvrages à proximité. Il s'agit dans cette étude de caractériser ces coefficients à partir de mesures de débit associées à des mesures de hauteurs d'eau. Un logiciel de calcul hydraulique donne les hauteurs d'eau à partir des coefficients de Strickler. Des algorithmes d'inversion probabiliste sont mis en œuvre afin de modéliser la densité de probabilité des coefficients de Strickler. Deux approches sont testées : l'une est une variante de l'algorithme EM, l'autre est de type MCMC.

Abstract: The statistical quantification of the roughness coefficients of the bed and sides of a river (minor and major Strickler's coefficients) enables both to assess the risks of an inundation and to reasonably size the nearby industrial facilities. The present work aims at inferring those coefficients from measures of the flow rate associated with ones of the water level. A hydraulic software is available to compute the water level in function of the flow rate and of the Strickler's coefficients. Two probabilistic inverse algorithms are applied to estimate the distribution of the latter. They corresponds to two different statistical approaches: a Maximum Likelyhood one by a variant of EM (ECME) and a bayesian one using MCMC.

1 Le problème statistique

Le problème motivant cette étude est celui de la modélisation hydraulique d'une portion de la Garonne. Bien que cette portion ne présente pas d'installation industrielle particulière, elle constitue un bon cas test à la fois pour les études hydrauliques [1] et les analyses de risque, de par les similarités qu'elle présente avec des portions fluviales à forts enjeux (notamment la sûreté nucléaire) pour EDF. Il s'agit d'une portion de la Garonne d'environ 50 km comprise entre Tonneins (Lot et Garonne) et La Réole (Gironde) ; voir la figure 1. Plus précisément, la question est celle de la quantification de la loi de probabilité de certains paramètres des équations hydrauliques, appelés coefficients de Strickler et supposés évoluer aléatoirement, sur la base d'observations indirectes, en fait des mesures de hauteurs d'eau pour certains débits.

Les équations physiques considérées qui lient ces grandeurs sont issues des équations de Barré de Saint Venant 1D, qui régissent l'évolution hydraulique du fleuve et font intervenir une abscisse curviligne x , le temps t , la hauteur d'eau z , le débit Q , la section mouillée S , les apports de débits latéraux q_L , la pente du tronçon i et les pertes de charges par frottement j entre l'eau et le lit fluvial. Pour les finalités de cette étude, nous supposons que le débit de crue est constant (cas stationnaire) et que les apports latéraux sont nuls :

$$\frac{\partial Q}{\partial t} + \frac{\partial(Q^2/S)}{\partial x} + gS \frac{\partial z}{\partial x} = gS(i - j).$$

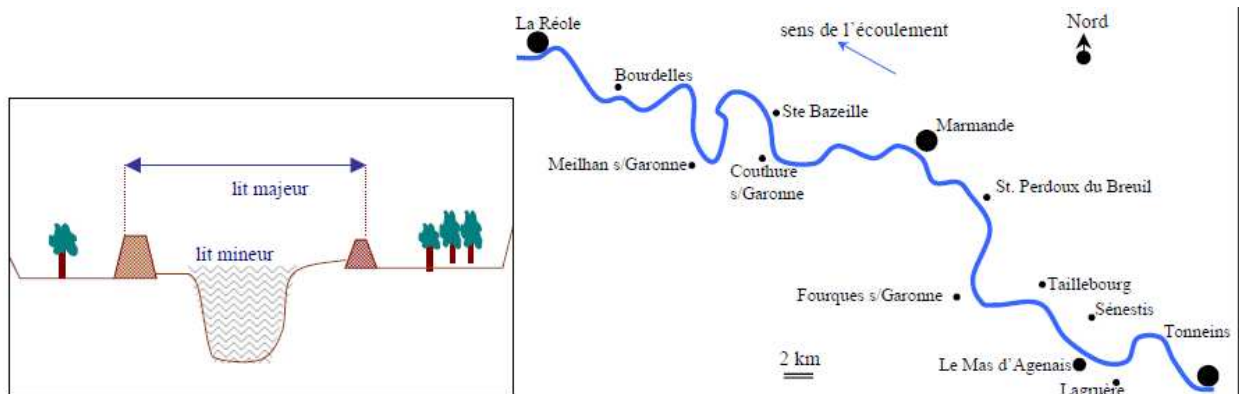


Figure 1: Configuration du tronçon fluvial.

Une modélisation dite "en lit composé" est adoptée, où chaque profil de la section fluviale (identifié par une valeur de l'abscisse x) comporte deux zones : un lit mineur (zone principale d'écoulement) et un lit majeur (zone élargie qui est investie en présence de crues importantes) ; voir la figure 1. Ces deux zones étant de nature différente, il leur est associé des rugosités différentes, ce qui se traduit par deux coefficients de Strickler, Ks_1 et Ks_2 . Ces coefficients permettent d'évaluer, en fonction du débit et de la morphologie de la section fluviale, les pertes de charges j . Le calcul hydraulique est réalisé à l'aide du logiciel "Mascaret" développé par EDF-R&D et le CETMEF (Centre d'Etudes Techniques Maritimes et Fluviales).

Notons G le résultat donné par Mascaret : $z(x) = G(Q, \mathbf{Ks}, \mathbf{d})$. Dans cette formulation, utilisant un formalisme répandu dans la pratique industrielle [2], nous avons isolé les grandeurs physiques qui sont connues avec incertitude : il s'agit du débit de crue Q , par nature aléatoire, et des coefficients de frottement \mathbf{Ks} , affectés par une variabilité temporelle supposée aléatoire.

Cet article décrit les résultats obtenus sur la caractérisation des coefficients \mathbf{Ks} obtenus par inversion probabiliste. Ces outils ont été utilisés sur un ensemble de $n = 100$ observations effectuées en deux points particuliers le long de la rivière (Mas d'Agenais et Marmande), une observation étant la mesure conjointe du débit ($Q_i \in \mathbb{R}, i = 1, \dots, n$) et de la hauteur d'eau en ces deux points ($z_i \in \mathbb{R}^2, i = 1, \dots, n$).

Les hypothèses et les choix considérés :

- H1 : Le fleuve est constitué de plusieurs tronçons. Le lit du fleuve est supposé homogène sur chaque tronçon (\mathbf{Ks} constant) et on suppose que seuls les coefficients Ks_1 et Ks_2 du dernier tronçon sont à caractériser. Leur densité de probabilité est supposée gaussienne $\mathcal{N}(\mu, \Sigma)$, la matrice de covariance Σ étant représentée par les deux écarts-types (σ_1, σ_2) et par le coefficient de corrélation ρ . Au total, 5 paramètres $\beta = (\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$ sont donc à estimer. On note $f_{\mathbf{Ks}, \beta}$ la densité gaussienne correspondant aux paramètres β .
- H2 : L'erreur ϵ entre les hauteurs observées z et le modèle est supposée être un bruit blanc additif gaussien, centré et d'écart-type $\sigma_\epsilon = 1\text{cm}$. On note f_ϵ la densité de la

loi gaussienne $N(0, \sigma_\epsilon^2 I_2)$.

- H3 : on dispose éventuellement d'un a priori sur $\beta : \pi(\beta)$.

La vraisemblance du vecteur de paramètres β s'écrit :

$$L(\mathbf{z}; \beta) = \prod_{i=1}^n L_i(z_i; \beta) \text{ avec } L_i(z_i; \beta) \propto \pi(\beta) \iint_{\mathbf{Ks}} f_{Ks, \beta}(\mathbf{Ks}) f_\epsilon(z_i - G(Q_i, \mathbf{Ks}, d_i)) d\mathbf{Ks} \quad (1)$$

L'utilisation de cette vraisemblance est rendue difficile par les expressions intégrales. Les approches utilisées visent à simplifier cette expression afin de pouvoir soit la représenter (approche de type MCMC) soit la maximiser (approche de type EM).

2 Approche du type EM (linéarisation itérée et ECME)

Le problème de la caractérisation des coefficients de Strickler peut être perçu comme un problème d'inférence statistique à données manquantes (qui correspondent alors aux réalisations non observées des coefficients de Strickler), c'est-à-dire qu'il peut être aisément réexprimé dans le formalisme proposé par Dempster *et al.* pour l'approche EM [3], où l'objectif est de maximiser la vraisemblance (dite incomplète) des observations $L(\mathbf{z}; \beta)$ dans une optique non bayésienne où l'*a priori* sur β n'est pas pris en compte : $\pi(\beta) = 1$.

Notons qu'une référence importante pour ce qui est du traitement de "problèmes inverses" par des méthodes de type (*Stochastic*) EM est le mémoire de thèse d'Estelle Kuhn [4].

Dans cette section, nous nous proposons d'illustrer l'application de cette approche, en fait de la méthode ECME [5], pour l'estimation de β . Cet algorithme est une alternative intéressante à EM, puisqu'il converge souvent plus rapidement que EM dans des situations où celui-ci s'avère lent. Plus précisément, nous allons exploiter les résultats de Celeux *et al.* [6], qui proposent, en outre, de combiner une linéarisation itérée de la fonction G avec un algorithme ECME pour estimer les paramètres d'un modèle statistique similaire. Dans la suite, nous nous limitons à une description détaillée de l'algorithme tel qu'il a été implémenté, permettant ainsi au lecteur de le reproduire aisément : pour une présentation plus générale de ECME ou sa déclinaison pour le modèle statistique considéré, nous renvoyons le lecteur aux références précédentes.

L'algorithme que nous avons mis en place est le suivant :

1. Choix d'un β initial : β^0 . On pose $k = 0$.
2. **Linéarisation** de la fonction G en $\mathbf{Ks}^k = (\beta_1^k, \beta_2^k) = (\mu_1^k, \mu_2^k)$, pour chaque observation Q_i ($i = 1..n$) ; en fait évaluation des $G_i^k = G(Q_i, \mathbf{Ks}^k, \mathbf{d})$ et jacobienes $J_i^k = \partial_{\mathbf{Ks}} G(Q_i, \mathbf{Ks}^k, \mathbf{d})$ de $G(Q_i, ., \mathbf{d})$ au point \mathbf{Ks}^k .
3. **Application de m itérations ECME** afin de maximiser la vraisemblance $L(\mathbf{z}; \beta)$ où G a été remplacée par son approximation de Taylor du premier ordre en $(Q_i, \mathbf{Ks}^k, \mathbf{d})$:

- (a) On pose $j = 0$, $\beta^{k,0} = \beta^k$, $\mu^{k,0} = (\beta_1^{k,0} \ \beta_2^{k,0})^\top$ et $\Sigma^{k,0} = \begin{pmatrix} (\beta_1^{k,0})^2 & \beta_1^{k,0} \beta_2^{k,0} \beta_3^{k,0} \\ \beta_1^{k,0} \beta_2^{k,0} \beta_3^{k,0} & (\beta_2^{k,0})^2 \end{pmatrix}$
et, pour $i = 1..n$, $\Delta z_i^{k,0}$, $M_i^{k,0}$, $C_i^{k,0}$ et $\mathcal{L}^{k,0}$ suivant les expressions données aux étapes 3e, 3c et 3f (avec $j = -1$).
- (b) $\Sigma^{k,j+1} = \Sigma^{k,j} + \frac{1}{n} \sum_{i=1}^n \left[\left(M_i^{k,j} C_i^{k,j} \Delta z_i^{k,j} \right) \left(M_i^{k,j} C_i^{k,j} \Delta z_i^{k,j} \right)^\top - M_i^{k,j} C_i^{k,j} M_i^{k,j}{}^\top \right]$
- (c) Pour $i = 1..n$: $M_i^{k,j+1} = \Sigma^{k,j+1} J_i^{k\top}$ et $C_i^{k,j+1} = (R + J_i^k M^{k,j+1})^{-1}$
- (d) $\mu^{k,j+1} = \mu^k + \left(\sum_{i=1}^n J_i^{k\top} C_i^{k,j+1} J_i^k \right)^{-1} \left(\sum_{i=1}^n J_i^{k\top} C_i^{k,j+1} \Delta z_i^{k,0} \right)$
- (e) Pour $i = 1..n$: $\Delta z_i^{k,j+1} = z_i - G_i^k - J_i^k (\mu^{k,j+1} - \mu^k)$
- (f) **Évaluation de la log-vraisemblance incomplète** (où la dimension p des observations vaut 2 en l'occurrence)¹ :
- $$\mathcal{L}^{k,j+1} = -\frac{1}{2} \left[n p \ln(2\pi) + \sum_{i=1}^n \left(-\ln(\det(C_i^{k,j+1})) + \Delta z_i^{k,j+1\top} C_i^{k,j+1} \Delta z_i^{k,j+1} \right) \right]$$
- (g) $j = j + 1$ et retour en 3b si $j < m$.

4. $\beta^{k+1} = \beta^{k,m}$ et $R = \sqrt{(\|\mu^{k,m} - \mu^{k,0}\|_2^2 + \alpha \|\Sigma^{k,m} - \Sigma^{k,0}\|_2^2) / (\|\mu^{k,m}\|_2^2 + \alpha \|\Sigma^{k,m}\|_2^2)}$.

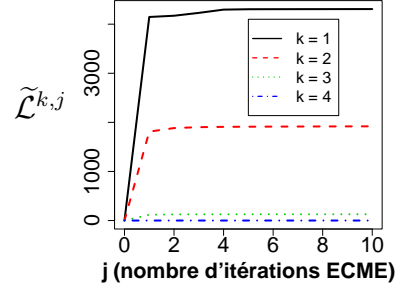
5. Si $R \geq \delta$, alors $k = k + 1$ et retour en 2, sinon $\hat{\beta} = \beta^{k+1}$.

Afin d'aboutir à une implémentation efficace, les $C_i^{k,j}$ ne sont jamais évalués et le calcul des termes de la forme $A C_i^{k,j} B$ est réalisé en trois étapes : (i) calcul de $L_i^{k,j}$ triangulaire inférieure telle que $L_i^{k,j} L_i^{k,j\top} = R + J_i^k M^{k,j}$ par factorisation de Cholesky ; (ii) calcul de \tilde{A} et \tilde{B} tels que $L_i^{j,k} \tilde{A} = A^\top$ et $L_i^{j,k} \tilde{B} = B$ par remontée ; (iii) calcul de $\tilde{A}^\top \tilde{B}$. Pour certains termes, $A = B^\top$: une seule remontée est alors nécessaire. Par ailleurs, les facteurs de Cholesky $L_i^{k,j}$ ne sont calculés qu'une seule fois et le calcul de la vraisemblance est simplifié, puisque le produit des termes diagonaux de $L_i^{k,j}$ est égal à $(\det(C_i^{k,j}))^{-\frac{1}{2}}$.

Le critère d'arrêt $R < \delta$ assure que l'algorithme stoppe lorsqu'une nouvelle linéarisation n'est pas pertinente, dans la mesure où alors, quel que soit α , le nouveau point de linéarisation serait identique au précédent à moins de δ près. En outre, pour $\alpha \neq 0$, il permet de tester la convergence de l'estimation de β . Néanmoins, le critère d'arrêt $j = m$ pourrait être modifié afin de tenir compte de la vraisemblance incomplète ou de son gradient. Il faut cependant noter que l'on peut choisir m grand pour des applications comme la nôtre de dimension petite et pour lesquelles l'évaluation de la fonction G ou de sa jacobienne constitue le principal coût informatique ; en outre, même pour m petit, l'algorithme ne stopperait que si l'estimation de β avait convergé au sens de $R < \delta$.

¹Comme l'expression de la log-vraisemblance le montre, $\Delta z_i^{k,j+1}$ et $(C_i^{k,j+1})^{-1}$ correspondent respectivement à $z_i - \mathbb{E}[Z_i]$ et $\text{Cov}[Z_i]$ où Z_i est une variable aléatoire de densité $z \mapsto L_i(z; \beta^{k,j+1})$.

Les “paramètres” de l’algorithme précédent sont donc β^0 , m , α et δ ; pour notre application nous avons choisi les valeurs suivantes : $\beta^0 = (10, 10, 1, 1, 0.5)$, $m = 100$, $\alpha = 1$ et $\delta = 0.01$. Par ailleurs, les jacobiennes sont obtenues par différences finies centrées (ordre 2) pour une perturbation relative de l’ordre de 10^{-6} . Quatre itérations suffisent pour que l’algorithme converge (voir la figure ci-contre). Les paramètres estimés sont les suivants : $\hat{\beta} \approx (32.84, 12.03, 0.84, 1.33, 0.32)$, ce qui est physiquement acceptable.



Log-vraisemblances incomplètes normalisées $\tilde{\mathcal{L}}^{k,j} = \mathcal{L}^{k,j} - \mathcal{L}^{k,0}$ pour $j \leq 10$.

3 Approche de type MCMC

La densité de probabilité de l’équation (1) ne peut être simulée directement par MCMC. Nous l’avons approximée en exploitant la faible dispersion des erreurs ϵ : pour des variations de la fonction G de l’ordre de σ_ϵ , cette fonction peut être approximée linéairement autour d’un point de fonctionnement $\overline{\mathbf{K}\mathbf{s}}$: $G(Q, \mathbf{K}\mathbf{s}, \mathbf{d}) = G(Q, \overline{\mathbf{K}\mathbf{s}}, \mathbf{d}) + J(\mathbf{K}\mathbf{s} - \overline{\mathbf{K}\mathbf{s}})$ où comme précédemment, $J = \partial_{\mathbf{K}\mathbf{s}} G(Q, \mathbf{K}\mathbf{s}, \mathbf{d})$ (matrice 2×2). Les z_i ont été observés pour des valeurs de $\mathbf{K}\mathbf{s}$ différentes et il serait logique d’utiliser un point de fonctionnement $\overline{\mathbf{K}\mathbf{s}}$ propre à chaque observation. Une stratégie pourrait être de considérer comme point de fonctionnement, pour l’observation i , le jeu de paramètres maximisant la vraisemblance de cette observation : $L(z_i; \mathbf{K}\mathbf{s}) \propto f_\epsilon(z_i - G(Q_i, \mathbf{K}\mathbf{s}, \mathbf{d}_i)) \pi(\mathbf{K}\mathbf{s})$, $\pi(\mathbf{K}\mathbf{s})$ reflétant l’a priori imposé à β . Ici, un a priori uniforme a été considéré pour chacune des composantes du vecteur β .

Pour l’application considérée, la dispersion des valeurs de $\mathbf{K}\mathbf{s}$ est relativement faible et un point de fonctionnement unique suffit. Nous choisissons pour ce point les valeurs de $\mathbf{K}\mathbf{s}$ permettant de maximiser la vraisemblance de tout l’échantillon, soit une fonction proportionnelle à $\prod_{i=1}^n L(z_i; \mathbf{K}\mathbf{s})$. Cette maximisation est effectuée en gardant le mode d’une simulation de type MCMC obtenue à l’aide d’un algorithme de type Metropolis Hastings (voir [7] par exemple). Les valeurs retenues comme point de fonctionnement sont $\overline{\mathbf{K}\mathbf{s}} = (32.75, 12.27)$.

Utilisant l’approximation linéaire dans l’équation (1), il vient :

$$L(z_i; \beta) \propto \pi(\beta) \frac{\sqrt{\det(\Sigma'_i)}}{\sqrt{\det(\Sigma)}} \exp \left(-\frac{1}{2} \mu^\top \Sigma^{-1} \mu + \frac{1}{2} v_i^\top \Sigma'_i v_i \right)$$

$$\text{où } (\Sigma'_i)^{-1} = \Sigma^{-1} + \frac{1}{\sigma_\epsilon^2} J_i^\top J_i \text{ et } v_i = \Sigma^{-1} \mu + \frac{1}{\sigma_\epsilon^2} J_i^\top (z_i - G(Q_i, \overline{\mathbf{K}\mathbf{s}}, \mathbf{d}_i) + J_i \overline{\mathbf{K}\mathbf{s}})$$

J_i étant le jacobien du modèle G au point $(Q_i, \overline{\mathbf{K}\mathbf{s}}, \mathbf{d}_i)$. Cette vraisemblance, connue à une constante près (la densité de \mathbf{z}), a été simulée à l’aide d’un algorithme de type Metropolis Hastings. Les histogrammes des coefficients du vecteur β (figure 2) montrent que les densités de ces coefficients sont vraisemblablement des fonctions régulières, voire unimodales. Les

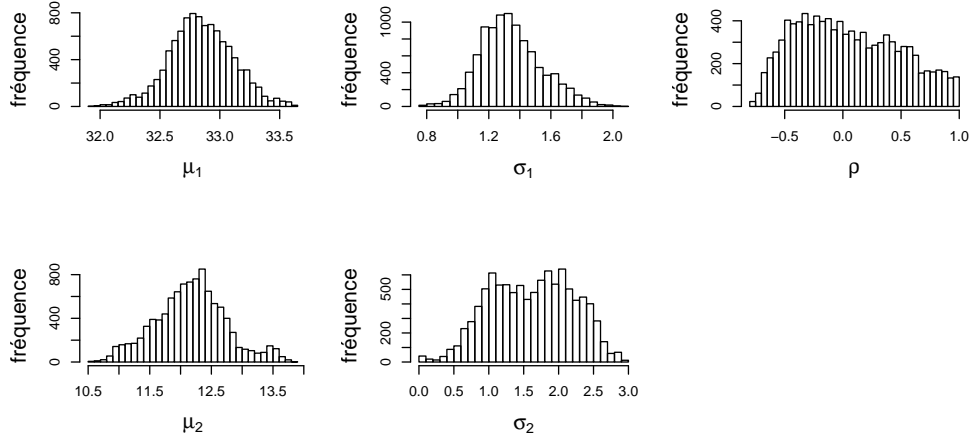


Figure 2: Histogramme des coefficients μ_1 , μ_2 , σ_1 , σ_2 et ρ simulés.

distributions obtenues sont compatibles avec les valeurs les plus probables identifiées avec ECME (section 2).

Remerciements. Cette étude a été en partie financée par l’Agence Nationale de la Recherche française (ANR) dans le contexte du projet OPUS (réf. ANR-07-CIS7-010 et ANR-07-TLOG-015). Par ailleurs, les auteurs remercient Nicole Goutal (EDF R&D/LNHE) pour son aide à la mise en place du cas d’étude.

References

- [1] A. Besnard et N. Goutal. Comparison between 1D and 2D models for hydraulic modeling on a flood plain: Case of Garonne river. In: *International Conference River Flow 2008, 3-5th September 2008, Cesme-Izmir (Turkey)*, 2008.
- [2] E. de Rocquigny, N. Devictor et S. Tarantola (eds.) *Uncertainty in industrial practice. A guide to quantitative uncertainty management*, J. Wiley & Sons, Chichester, 2008.
- [3] A. P. Dempster, N. M. Laird, D. B. Rubin. *Maximum Likelihood from Incomplete Data via the EM Algorithm*. Journal of the Royal Statistical Society, Series B, Vol. 39, No. 1, pp. 1-38, 2007.
- [4] E. Kuhn. *Estimation par maximum de vraisemblance dans des problèmes inverses non linéaires*. Mémoire de thèse, Université Paris XI d’Orsay, 2003.
- [5] C. Liu et D. B. Rubin. *The ECME algorithm: a simple extension of EM and ECM with faster monotone convergence*. Biometrika, Vol. 81, No. 4, pp. 633-648, 1994.
- [6] G. Celeux, A. Grimaud, Y. Lefebvre et E. de Rocquigny. *Identifying intrinsic variability in multivariate systems through linearised inverse methods*. Rapport de recherche INRIA n° RR-6400, <http://hal.inria.fr/inria-00200113/fr/>, 2007
- [7] C. P. Robert et G. Cassella. Monte Carlo Statistical Methods, Edition Springer, 2005.